O. Sami Saydjari, ssaydjari@cyberdefenseagency.com
Vijay Varadharajan, vijay@ics.mq.edu.au

The Evolution of Online Identity

Scott Charney
Microsoft

When I look out "On the Horizon" to think about emerging Internet trends, I think that as a society we are beginning to see changes that can improve how we manage our identities online. In large part, these changes are necessary because, to reduce online crime, we must significantly improve how we authenticate ourselves on various computer systems.

The Need for Improved Online Identity Systems

We need only juxtapose the physical and cyber worlds to see the need for improved online identity systems. In the physical world, identity is based on social custom, followed by the creation of identity documents and derivative identity documents. By way of example, a child will likely be named at birth, which is when his or her first identity document—the birth certificate—will be created. This document is later used to create additional public- and private-sector identity documents: when the child is ready to drive, he or she will produce that birth certificate to get a driver's license; when the child wants to open a bank account, the bank will use that driver's license to open an account and issue a bank card; when that child wants to travel overseas, the post office will ask for two forms of identity, the birth certificate and the driver's license (which, of course, was issued based on the birth certificate) before issuing a passport.

Compare this process to the one we use to create an Internet identity. This same person will go to a Web site and enter "secret" data (such as his birth date and mother's maiden name), the recipient Web site will verify this data with a third party, and identity will be established. The problem, of course, is that this "secret" data is not secret at all: many people have access to this information and could inappropriately use it to "authenticate" themselves as this person. It becomes clear, therefore, that if we want to apply solutions from the physical world to problems of identity on the Internet, we must create digital Internet identities that are based on in-person proofing (IPP) and the issuance of true secrets (digital certificates) that permit unique identification claims. With these tools in place, people can assert identity or, even better, identity attributes (such as age or residency), enabling other people and organizations to more safely trust that information. Of course, this identity system will not be perfect, but physical identity documents are subject to misuse, too. The point is that we can create online identity systems that are more robust than the ones in place today. Moreover, these online identity systems could provide greater protections for security and privacy than we currently achieve when using documents to prove identity in the physical world.

As I mentioned earlier, online identity management systems are important if we're to address the Internet's growing crime problem. The range of criminal activity that the Internet supports is broad, including consumer threats (such as compromised computers being used for unauthorized activities, identity theft, financial fraud, and child endangerment), enterprise threats (such as the theft of financial information, loss of personally identifiable information, economic espionage, and extortion via threats of denial-of-service attacks), and government threats (such as information warfare). These crimes are pervasive in part because the Internet has four attributes that make it attractive to criminals: global connectivity, anonymity, a lack of traceability, and valuable targets. Without proactive controls (such as neighborhood watches and police patrols) and absent reactive effectiveness (due to anonymity and lack of traceability), those who commit crimes on the Internet have little concern about identification and capture and, therefore, little to deter them. This is one reason why identity management is so important.

Not surprisingly, however, mentioning the words "identity" and "the Internet" in the same sentence gives many people pause, in large part because the Internet has been so transformative in the areas of free speech and communication—areas where anonymity plays an important part in ensuring the free flow of ideas. Social networking represents the new town square, and blogging has turned citizens into journalists. Therefore, while the thought of strong digital identities cannot be proposed lightly, absent a way to create, transmit, and consume robust identity on the Internet, people will lack the data necessary to protect their own security and privacy online. To enable robust identity on the Internet, we need to create an "identity metasystem" to enable better trust decisions and help solve difficult real-world problems such as identity theft.

The current strategy for addressing identity theft has two parts. The first prong is to educate consumers not to disclose personally identifiable information (PII) improvidently. The second prong requires that data custodians (that is, the people and organizations that hold PII) not lose it. Both of these strategies are important, but neither works on a large scale. Although many consumers might make good judgments about disclosing PII, and many companies engage in reasonable security practices and avoid data breaches, identity theft will remain a problem so long as a large number of consumers and some number of data holders lose sensitive data. Thus, it becomes clear that the key to combating identity theft is to devalue PII. If individuals were given a true secret (such as a certificate) based on IPP, then the reliance on "shared secrets" would be eliminated. By reducing the use of shared secrets, cybercriminals would no longer have access to the key pieces of information they need to consummate a fraudulent transaction. For example, a cybercriminal might apply for a loan using another person's "shared secret" data (such as birth date and mother's maiden name). If the financial institution required applicants to enter a smart card and PIN to confirm their identities before a transaction is consummated, cybercriminals would be thwarted. (A larger discussion of these issues appears at www.microsoft.com/endtoendtrust and www.microsoft.com/endtoendtrust/lwsd.)

Addressing Anonymity and Privacy Concerns

Although necessary and beneficial, the creation of an identity metasystem raises important social issues. Two of the more pressing concerns relate to protecting anonymity and privacy. The first concern is that if authenticated identity is required to engage in Internet activity, anonymity and the benefits that it provides will be reduced. Although anonymity might exist on the Internet due to historical evolution, the fact is that it serves many useful purposes. For example, anonymity supports important policies regarding the promotion of free speech, even if harm sometimes occurs because of the anonymous nature of the communication. Indeed, it is important to remember that some societies have long accepted and promoted anonymous speech, despite these concerns. This is why it is still possible to make anonymous phone calls (pay phones being replaced with disposable cell phones), and you can mail packages (containing contraband) without a return address. Even with the potential risks that anonymous Internet speech can bring, there are both practical and philosophical reasons to continue to permit it.

Although necessary and beneficial, the creation of an identity metasystem raises important social issues. Two of the more pressing concerns relate to protecting anonymity and privacy.

That said, it is an overstatement to say that the Internet, although imbued with various types of communication, is simply about communication, or that it is akin to other forms of communication networks. For example, the Internet provides communication abilities on a scale previously unknown. Although you could in theory call or send paper mail to millions of victims, the time and cost involved make doing so infeasible. Moreover, the Internet's multipurpose nature and power make a comparison between it and other communication technologies overly simplistic. Comparisons to traditional voice networks fail because the Internet is not just—or even primarily—about voice. Cybercriminals can attack critical infrastructures in a way phone users can't. But because many people access the Internet from private places such as their homes, permitting greater attribution of activities might be worrisome to some.

The second concern is that authenticated identifiers could be aggregated and analyzed, thus facilitating profiling (although there is certainly concern about data profiling even in the absence of an identity metasystem). Three factors, however, help mitigate this concern. First, people will have many forms of identity and can provide different identifiers in different contexts, thus reducing the risk of profiling. Even today, people have many personas (work identities, personal identities, pseudonyms, and temporary or anonymous "identities") and many identity documents (a state ID, a federal ID, a bank ID, and an employer ID). In the context of an identity metasystem, each user should be able to choose what identity to use in a given situation.

Online identity systems could provide greater protections for security and privacy than we currently achieve when using documents to provide identity in the physical world.

Second, the use of identity attributes, as opposed to sharing your full identity, should help protect privacy. For example, if you want to visit sites with content not appropriate for children, you should be able to prove your age without necessarily providing your name. In this sense, an online identity metasystem is more privacy protecting than offline systems. When I was younger, certain establishments would ask for my driver's license to verify that I was old enough to drink alcohol. When I handed them my license, they were interested in just two data points: my picture (to make sure it was indeed my license) and my birth date (to make sure I was old enough to imbibe legally). Still, I had to provide my entire license, which contained other PII that the proprietor had no need to know, such as my name and address. Focusing on specific personal attributes could even enable new, privacy-centric business models. For example, it might be possible to engage in targeted anonymous advertising because it is increasingly possible to "know" something about someone's interests without knowing who they are. This is not beneficial only for merchants and advertisers who want a greater return on their investments but also for consumers because many free services, such as email and search, are actually paid for by advertising revenue.

Finally, social rules can be constructed to support anonymity in appropriate contexts. Governments, for example, could require the use of robust identification when someone claims government benefits but prohibit demands for identification when someone seeks public information from a government Web site.

Clearly, this approach might not satisfy those who see the Internet's anonymity as the ultimate protector of privacy and an identity metasystem as a threat to greater anonymity. The fact remains, however, that if we hope to reduce crime and protect privacy, we need to give users the ability to know with whom they are dealing (if they so choose) and give law enforcement the capability to find bad actors. It is also important to remember that multiple privacy interests are at stake. For example, in the email context, it is not just a communication's sender who might have a privacy interest: the recipient might also wish to be left alone. Indeed, any regime should not only seek to provide greater authentication to those who want to provide or consume it, but also provide anonymity for those who wish to engage in anonymous activities. Users should be able to choose to send anonymous communications or receive mail only from known sources. Users who want to accept anonymous communications should be free to do so, but they should also understand that they might have little recourse if that anonymous communication proves to be harmful (such as a threat or fraud). The bigger "philosophical" issue relates to the fear that if an authenticated infrastructure is available, then neither market nor social forces will support a vibrant anonymous culture. Put another way, if authentication were possible, what if every social networking site, email system, and Web site required authenticated identities? How would the social values promoted by anonymity be supported?

Although this debate cannot be resolved to everyone's satisfaction because it is impossible to prove what will happen a priori, we could argue that people have long shown an interest in and support for anonymity; markets will support anonymity, much as you can shop today without providing proof of identity; and anonymity and privacy protections can be established through regulation.

The Future: Creating an Online Identity Metasystem

Given these arguments, if we agree that an identity metasystem's benefits outweigh its risks, the challenge is to create this IPP-based identity metasystem. Such a system requires five components.

First, for consumers to obtain robust digital credentials, we need organizations capable of conducting IPP. The IPP locations must be ubiquitous but can be either public or private institutions. For example, public (or quasi-public) institutions that currently engage in IPP activities include the Department of Motor Vehicles, which issues not just driver's licenses but identification cards; post offices, which proof identities for passports; schools, which enroll students based on IPP events (with children and often their parents present); and financial institutions, which use documents such as driver's licenses to issue derivative documents with identification information (such as credits cards).

Second, we need organizations to manage identity claims, including revoking certificates when credentials are lost. In some cases, the IPP entity might also issue and manage the IT infrastructure necessary to transmit claims and revoke certificates. In other cases, however, the organization that conducts the

IPP event and the organization that issues, manages, and revokes digital certificates might be different. For example, post offices conduct IPP events for passport generation, but the US Department of State actually issues the passport.

Third, we need easy-to-use formats that are supported by widely available technology. For example, magnetic stripes are familiar to consumers, and the security issues associated with such technology might not be problematic if the only data encoded on the stripe is meant to be public (such as data signed with a private key that is meant to be shared and then verified with a public key). Smart cards allow for computations, but neither smart cards nor card readers are currently ubiquitous, particularly in the consumer space. Other forms of two-factor authentication might include USB dongles and smart phones.

Fourth, we need to ensure social, political, economic, and information technology alignment. For example, at the same time consumers obtain such certificates, governments and businesses must build the infrastructure necessary to consume such identities, and policy makers must create a regulatory framework that advances—or at least does not inhibit—the identity metasystem. Many years ago, a hardware vendor showed me a consumer keyboard with a magnetic stripe reader. You could quickly see its value in reducing online credit-card fraud since such technology might enable a consumer to prove to a merchant that the consumer actually had the credit card he or she was using. (Yes, some criminals can duplicate magnetic stripes.) When I asked the hardware vendor about the sale of this keyboard, he noted that consumers saw no reason to pay the slight differential required. Not only were consumers very sensitive about price, but, more importantly, no Web site ever asked them to swipe their card.

I then consulted with a security professional at a bank. When I suggested that such keyboard readers might reduce credit-card fraud, he noted that this might actually increase the bank's financial risk. The reason, he said, was that there were two types of credit-card transactions: a card-present transaction (in which the merchant verifies that the consumer has the card) and a card-not-present transaction (such as a purchase over the Internet or via a toll-free number). He noted that in a card-present transaction, the merchant bank pays the merchant even if the credit card turns out to be fraudulent. In contrast, the merchant takes the loss in a card-not-present transaction. This being true, and recognizing that the magnetic stripe reader might turn a card-not-present transaction into a card-present transaction, he noted that such keyboards might shift the risk of loss from the merchant to the bank.

So, if this keyboard were to help merchants and not banks, it was time to talk to merchants. Although those I talked to were certainly interested in fraud reduction, they noted that it would take considerable work to build the back-end infrastructure necessary to enable this card swipe technology at the consumer level. There was no point making this investment, they noted, because no consumers had keyboards with magnetic card readers.

It was a classic chicken-and-egg market failure: consumers would not pay for a keyboard reader because Web sites did not ask them to swipe their credit cards; Web sites did not ask consumers to swipe their cards because they knew consumers did not have keyboard readers. One way to address market failure is, of course, with government intervention. Noting that the Office of the Comptroller of the Currency put in place regulations promoting the use of two-factor authentication at banks, I asked government personnel whether engagement on this chicken-and-egg problem was likely. They said no—such action would constitute interference in the market, which was true, because it might drive the market toward a particular solution. Lesson learned; without alignment, success is not possible.

Fifth, it must be remembered that criminals are creative, adaptive, and persistent. Therefore, any identity metasystem must have a carefully constructed and comprehensive threat model. Although robust digital identities based on IPP and digital certificates might make it harder for criminals to impersonate others and commit crimes, we should expect that criminals will find new ways to circumvent these defensive measures. For example, a criminal might bribe an IPP agent, steal a valid certificate and PIN, steal the keys used to sign certificates, or social engineer a call center after claiming to have lost a digital certificate. These and other threats should be considered and mitigated by business process and technology.

If we want the Internet to reach its full potential, we need a safer, more trusted online environment. To achieve this, we at Microsoft have proposed a vision outlining the reasons for end-to-end trust. But Microsoft and the technology industry alone can't create a trusted online experience. For this to happen, industry must not only band together but work with customers, partners, governments, and security and privacy experts worldwide to help take trustworthy computing to the Internet.

Scott Charney *is vice president of trustworthy computing at Microsoft. His research interests include security, privacy, and public policy regarding those issues. Contact him at scharney@microsoft.com.*

Top articles, podcasts, and more.
cn
computingnow.computer.org